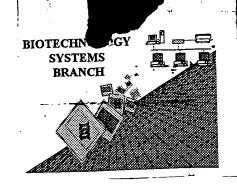
RAW SEQUENCE LISTING
ERROR REPORT

Date Processed by STIC:



13 Sunda 8/18/00

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/10/423Source: 1632

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR FURTHER INFORMATION, PLEASE TELEPHONE MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER</u> <u>VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address: http://www.uspto.gov/web/offices/pac/checker

•		SUCCESTED CORRECTION SERIAL NUMBER: 09/101,423
\ 1	ERROR DETECTED	SUGGESTED CONTROL
ATTN: !	Wrapped Nucleics	EASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it.
		Please adjust your right margin to .3, as this will prevent "wrapping".
2		The amino acid number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".
3		The rules require that a line not exceed 72 characters in length. This includes space FECH CENTER 1600/2900
	Numbering	The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
5	Non-ASCII	This file was not saved in ASCII (DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
6	Variable Length	Sequence(s) contain n's or Xaa's which represented more than one residue. As per the rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
7	Patentin ver. 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
		sections for Artificial or Unknown sequences.
8	Skipped Sequences (OLD RULES)	Sequence(s) missing. If intentional, please use the following format for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS") (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: This sequence is intentionally skipped
		Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
9	Skipped Sequences (NEW RULES)	Sequence(s) missing. If intentional, please use the following format for each skipped sequence. <210> sequence id number <400> sequence id number 000
10	Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of <220> to <223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
11 \(\sum_{-11} \)	Use of <213>Organism (NEW RULES)	2 missing this mandatory field or its response.
12	Use of <220>Feature (NEW RULES)	Sequence(s) are missing the <220>Feature and associated headings. Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown" Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rule)
13	Patentln ver. 2.0 "bug"	(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29031-02) (Sec. 1.028 at No. 104, pp. 29031-02) (Sec

AUG 16 WIII

1632

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/101,423

DATE: 08/09TECHCENTER 1600/2900 TIME: 14:53:45

Ivan thoughout.

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\1101423.raw

3 <110> APPLICANT: Rudland, Philip S.

Barraclough, Roger B. W--> 5 <120> TITLE OF INVENTION: Metastasis Inducing DNA=s

6 <130> FILE REFERENCE: WPT 0114 PUS

7 <140> CURRENT APPLICATION NUMBER: US 09/101,423

8 <141> CURRENT FILING DATE: 1998-11-27

9 <150> PRIOR APPLICATION NUMBER: PCT/GB97/00074

10 <151> PRIOR FILING DATE: 1997-01-10

W--> 11 <160> NUMBER OF SEQ ID: 6

Does Not Comply Corrected Diskette Needec

sel dem 5 on Eva Sumary

Meet

ERRORED SEQUENCES

13 <210> SEQ ID NO: 1 14 < 211 > LENGTH: 1033 base pairs

15 <212> TYPE: DNA

16 <213> ORGANISM: Homo sapiens

18 <400> SEQUENCE: 1

E--> 19 cttccttggt gctctatgtc ttgcctctcc ccttctccag tcccattaag 60

E--> 22 cttgacagac tctgggacag tcccctctgc tctcctgttg gcgcctgagt 120

E--> 25 ctgaggaccc ttcacgtagc ctcccatctg gatgacctag tagaagacgt 180 26 gggaagttgt

taactgagca gagctcagag atttaaagtg agtctgggga E--> 28 cacactcagg 240 29 gcctcgagga

gccttaaaaa gccaattgga tgactaaccc agactattgt E--> 32 ttgatctgct 33 cactttaggt 300

E--> 35 gggaagtcac tagcatatet gatgggtcac atctgagaaa ggtttetage 360 E--> 38 ttgtgtgage ageatggegt gtateatggt gtgcageata etcaggetge

420

E--> 41 cgaggetett etteagtatt aggggaacea etggtgttga acatggteea 480 E--> 44 catgtgagga gaatcccaat gcgtcaggag aaaacgagag tctgtgacct

540 E--> 47 agatacagaa ttattcttgg actgtgtttt catgctcctt gtggatggga

E--> 50 ttcaggttaa tcagcattgc ttactgttgg tattcaagta aatgcttaaa 600

660

E--> 53 tatacctctg tgggaagcag gtttttgata catgcagctt gtccttgtga 720

E--> 56 ttgaactcaa gagaactttg ctcatgtgat ctttcttaac cgatggagta

E--> 59 gatgetetea ataaagttgg etettgeacg agacgttagt etgteetgtt 57 gaaactgtct

E--> 63 attetteege teccaeggee tetacageae taaacecaee acegatagae

Jummany Steet

fless ensure Lover-case letters, per new Sequene Rubs format.

Lover-case.

PATENT APPLICATION: US/09/101,423

DATE: 08/09/2000 TIME: 14:53:45

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\I101423.raw

E--> 66 actgacaaac atcaccagag gctcttaact gagattataa actgttacta 900 E--> 69 ggaatcgctc cccagaaaca taaacattta cttggagaac tcaagacccc 1020 70 tttgtagaca

72 taactcccat ggt 74 <210> SEQ ID NO: 2

75 <211> LENGTH: 1058 base pairs

76 <212> TYPE: DNA W--> 78 (213> ORGANISM:

78 <400> SEQUENCE: 2 E--> 79 attgctgtga gcctattagc gacatttggt gacgcccctt ttaagggggt 60

E--> 82 aatgggttga aattetgtge cacaaacget etceatgttt teacaattae

E--> 85 tgtggtcagc agccagaatt tagggatgtg atggggacagg gtcggggaaa

E--> 88 ggtaaaggaa agacagcacg ttaaagtcca aacagctcca ggagactatc 240

E--> 91 acatcagacc atgaggagaa ttgatatcat tgtttttcaa tgggtatcgc 300

E--> 95 ttccatctga ttaaaaataa ttactgctgg cactaaatcc aattggaaat 360

E--> 98 atttatette caetteatge tgetaceata tgeetgaegt ggeggageag 420

E--> 101 tcccgttctg ataaatagta ctttgtaaat atttggagac gggagetctg 480

E--> 104 acacgtacaa accggcctgt ttatcatgtt cccgatagag gccctctttg

E--> 107 coccaaaaca gtcaggatgc tgtgaatttc cttccatgaa gccttgttca

E--> 110 ccattggagg aagcaggetg cactgtetac cacaagtgge actttecaaa

E--> 113 atattggagc aagacatttt gctggctgac tggtgctgtg taagctgata 720

E--> 116 tttattaaac tggcttttct ttgaacaccc cactcaagga aaaaaaaaca E--> 119 gacattattt ggagatgaag tetttataga gatgettaag tttaaacgag

E--> 122 ccggctctat tccatttaat gaatggtgtc cctacaaagg aagaaactgg

E--> 126 tgtacacttg tgtgtgtgt agagacaacg tgaggagctg aagaggagca 900

E--> 129 agagaaaggc tgacccttat tcacactgag caaaccagtc atgtgtgggt 1020

130 cgatagatga 132 gagtatecce caagacteae acattegaae gettggte

134 <210> SEQ ID NO: 3

135 <211> LENGTH: 1008 base pairs

136 <212> TYPE: DNA

A:\ES.txt
N:\CRF3\08092000\I101423.raw

A:\ES.txt
N:\CRF3\08092000\I101423.raw

A:\CRF3\08092000\I101423.raw

A:\CRF3\08092000

famet ever

1058

PATENT APPLICATION: US/09/101,423

DATE: 08/09/2000

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\I101423.raw

137 <213> ORGANISM: Homo sapiens

139 <400> SEQUENCE: 3

E--> 140 aggaccagag ttcacatccc atcacatggc ccagaaggtt ttaatgctgt 60

E--> 143 aggggggaac tgcacacac tgtgcacata cacttacaga gacacacatt 120 E--> 146 agaacacaat cacaaataaa aaaaatcttg aaaaatttta agctaaaatt

180

E--> 149 aacatatata caatttttct ttatttttt aaagatttat ttatttaatg

240

E--> 152 cactgootot cootocagac atagoagtac agggoatogg atoccattac 150 tatatgagta

300

E--> 155 gagccaccat gtggtttcac agatggttgt gagccaccat gtggtttcag

360

E--> 158 caggacettt ggaagageag teagtgetet taacetetaa gecatetete 156 gaattgaact

420

E--> 161 tatacaattt taatgctacg tacacacaac ttctctttcc tttaatggtt

480 E--> 164 tctggagaag taagaataaa ggagggaaag aacattgctt tcacattgca 162 gagatttttg

540 E--> 167 cagcgtgttt aaagtaggaa tgccatgaaa tgactggcct gccttctcat 165 ccagtgggaa

600 E--> 170 cccactcctc cttttaactg gageteettt atetaattta ttagtttgae

660 E--> 173 gttttettet gttttgatet ttttaagaca gagacteace atatageeet 171 gatacccagg

720

E--> 176 aagotcacta tytagaccag tetggcettg aactcaaagg agatetatet 780

E--> 179 gctgggatta aaggcttgtg ctaccaagtc tggtctgagg ctttggagca 840 180 gcctcggttt

E--> 182 tggccttctt taaggatete taagetagea gtaagtagee tagecatget E--> 185 gttgttcgtt catcotggct ccagcacaaa ggcagtcact aaacgtcggc 183 gttgtaggaa 900

186 ctcatttcat 960 E--> 189 cagagetgaa tgeaaattee ttgtgetett cetgtgteet eetggaae

E--> 190 1008

192 <210> SEQ ID NO: 4

193 <211> LENGTH: 1088 base pairs

194 <212> TYPE: DNA

195 <213> ORGANISM: Homo sapiens

197 <400> SEQUENCE: 4

E--> 198 agttggggac acagcttgct tgattaagat gtttcttggg aaaaggagtt

60

E--> 201 atttccaatg gaaaggactg ctaattgggg aggcaatgtt gcttaattgg 120 202 gacacctgcg

E--> 204 ggtaattaaa agetetetee cagtggeett teetgttttt ggetetggga 180

E--> 207 ttgagaggga tgcaggcatt ctaagggctg gttcttggtt tctcccttcc 205 ggcgaaggca 240 208 cctctgtcca

TIME: 14:53:45

RECEIVED

MR 18 SOON

TECH CENTER 1600/2900

PATENT APPLICATION: US/09/101,423

DATE: 08/09/2000 TIME: 14:53:45

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\I101423.raw

E--> 210 aactcagtga ggtatccctg tctgtgctgt ccttagagtg ccgtcctgag 300 E--> 213 gttaaggtet etggatetga getgeeteag ggaaacgeat gageteattg

E--> 216 aaccaggcaa aggtgttggc tgtgacctca gaattctgag gggcaaaggt 214 gaaaggggag 360

420 E--> 220 ctctcattat agagcaagtt tgagactggc ctgggaacaa aaatataaag 217 tcaaggctaa 480

E--> 223 catatgacag cacctgagga gtcctgtccc tagagatcat aaggacctgg

E--> 226 cttgttgcag atggcacttt gtgtcgagag aggggacctg ccccagcatg 224 ctgctgggga 600

E--> 229 gaagatcctc tggattaact gtgaacactg attgctgctt tatacctgga 660 230 gttgtgctgt

E--> 232 tatotggtac acatetgctg ggtgaatgag ttcatgggct ttatttcagt 233 gaggtattta 720

E--> 235 cctgaggaga aagaaggact ggtgccacaa agcacagctt ttaaatctgt 780 236 gggttgtgac

E--> 238 ccattatgga ctatcataac tgagtgcagg tatcaagaat actttagcag 239 gtggtaaaaa

E--> 241 gatttttgaa tgcgcaacga ccaaaactga actcaaaaat caagcatggc 900 242 atggatcctg

E--> 244 ggtgctcctg gaagcacttg cetttactgc attgtgcgac ttgacggtag 960 245 ccttggttct

E--> 247 gaatgcacaa cacgtgggct ttgggctgca caggccacca cgccgtgcct 1020 248 gaaacacctc

E--> 251 ageteaggtt tgtggetatg tectatgact tggaettact tttattgeac 1080 252 atataaatat

1088 254 tttcctgc

257 <210> SEQ ID NO: 5

258 <211> LENGTH: 960 base pairs

259 <212> TYPE: DNA

260 <213> ORGANISM: Homo sapiens

262 <400> SEQUENCE: 5

E--> 263 gagggggtgg tggcacagtt atgtttttgt aggaagggtt ccatgaacct 60 E--> 266 cgggttagaa atttaaaagc cctgagggga atttttttt taaatcgcta 264 cagcagagct

120 E--> 269 atgagaaaaa cagatcagaa acgttcttgt gcttcagaaa aggacaagtg 267 tgaatctgac

180 270 tgtgagctaa

E--> 272 cagactgcac actggtgtte gaggcacate tggatcacag gagcgtcaga 240 273 taatgtcccc

E--> 275 aaaggtaaat gcatttgctt gcacagtacc gagtgtggtg gggggtgcct 300 276 acagcccagc

E--> 278 ggttctcaac cttcctgatg cttcgaccct ttaatacagt gcctcatgct

360 E--> 281 ccccaacctt aaaattattt ttgttgctgt tcataactgt gattttgata 279 ctggtgacct

E--> 284 ttgtaatata aataattttg aagaaagagg tttgccaagg gtttgagaac

RECEIVED MR 18 SOOD

TECH CENTER 1600/2900

Money

PATENT APPLICATION: US/09/101,423

DATE: 08/09/2000 TIME: 14:53:45

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\1101423.raw

E--> 287 gccccacgtg gatggttttt cgtcatttgg ggtttttatg aggcagagtc 540 E--> 290 caggetagea geetagaatg tgetaettag etgaggaata acettggaae 288 ttatgtagcc 600 E--> 293 tggagagact ggcttagtcc tcaagaaact ggaaatagct ggagtttggc 291 ttctgaggac 660 ttcaaacett ttctactctt tttccaccct gtcggccccc 294 tacttgtggg E--> 296 ttcctttttc E--> 299 taagaaagag aaaggggagc atagagggga aaagaaaccc ctgaataacg 297 taacactaaa 300 tcagtagttg 780 gtgacatatg ttgtcattag accacatect ggtgattaag E--> 302 gcaaaggggg 840 303 gggagtcaag caagtttgat ctttcgtgta acgatatcta atttcttctc E--> 305 ttccttgggg E--> 308 cgtctttgtg aacaacgact tgataaccca caatggacca tcaaccaacc 306 cctgttgctt 900 960 309 aaccaaccat 311 <210> SEQ ID NO: 6 312 <211> LENGTH: 1090 base pairs 313 <212> TYPE: DNA 314 <213> ORGANISM: Homo sapiens 316 <400> SEQUENCE: 6 E--> 317 ttgtctctgg tgttacttgt tttcccattt ctgacagtgg tttgaccttc 60 E--> 320 tgtgtcagga gtgctgtaga cctattttcc tgttttcttt cagccagtta 120 E--> 323 gtgttctact gtcagatgtg tagctgttcc tgtccactga ctttcaagct 321 caggaacaga E--> 326 gcaggaacca gaagggcctg tecetaette tactgggccc etacgcacag 240 E--> 329 tggtgctagg tgttttcctc tagagcctga aatgtgggca gagagtagtc 327 ggggcctaga 300 E--> 332 tectaggtat gtetteeect etgaaggtet ageteteeet tecatgggat 360 E--> 335 ggagetgttt gaccaggtee teteaaatee gggtgeagte tggacegeag 333 atgggtgcag 420 336 gctcctgtag E--> 338 cttgcctgct gcaatcttcc cgcacccaga ggcacccaag tttcctcttg 480 339 ggccaaggat E--> 341 gtgggcaaag gtgggcagaa gtggcaatct ctcctgccct agcgtctcag E--> 344 acttetggge aatecgetet etettecaca gggtttggga geagggaget 342 gattgccctc 540 345 gtgggccggt 600 E--> 348 atcaggcaaa ggtttgaggc aaccagttag aaactggaag tgtcaggtcc 349 cagaggaatt E--> 351 ttgcctttgt gtgtcctgag tccaccaggc aggtcacttg gagcagaaaa E--> 354 ccctcggtct caggcctgaa gttgcacctc agggttggct ttcagctgta 780 E--> 357 gtatggtttt aaaaatctaa gatagctatc atgcagcaag gcttgtgtaa 355 cctgtggaaa

RECEIVED AUG 18 ZOOU TECH CENTER 1600/2900

same

PATENT APPLICATION: US/09/101,423

DATE: 08/09/2000 TIME: 14:53:45

RECEIVED

MIR JE SOON JECH CENTER 1600/2900

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\I101423.raw

358 aatgtctatt

E--> 360 tggttccttt atgacttact tttgctgtac tgaggatcaa acctagggtc 900

361 tcaagcagtc

E--> 363 atcacaattc totgtcactg atccagetcc atttctattt tottttgtcc 960 E--> 366 ctcgccagca agaaaacacg ctagggacat acgaatcctt gctgcagcca

1020

E--> 369 tgaatcttaa ggagaagccc gcgcaccgga ctggcgcggt ttatatacac 367 aaacttttat 1080

370 cctagcacag

W--> 374/M:\Clients\W\WPT\0114pusa\sequence list with pg #s.wpd E--> 375 sandy m. 7-21-00 E--> 376 shelly 07/24/00 372 tgcatccaca 1090

E--> 3/78 -14-

delete at end of file

VERIFICATION SUMMARY PATENT APPLICATION: US/09/101,423 DATE: 08/09/2000 TIME: 14:53:46

Input Set : A:\ES.txt

Output Set: N:\CRF3\08092000\I101423.raw

L:5 M:283 W: Missing Blank Line separator, <120> field identifier L:6 M:283 W: Missing Blank Line separator, <130> field identifier L:7 M:283 W: Missing Blank Line separator, <140> field identifier L:11 M:283 W: Missing Blank Line separator, <160> field identifier L:19 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:1 L:19 M:112 C: (48) String data converted to lower case, M:112 Repeated in SeqNo=1 L:78 M:282 W: Numeric Field Identifier Missing, <213> is required. M:254 Repeated in SeqNo=1 L:79 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:2 M:112 Repeated in SeqNo=2 M:254 Repeated in SeqNo=2 L:140 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:3 M:112 Repeated in SeqNo=3 M:254 Repeated in SeqNo=3 L:198 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:4 M:112 Repeated in SeqNo=4 M:254 Repeated in SeqNo=4 L:263 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:5 M:112 Repeated in SeqNo=5 M:254 Repeated in SeqNo=5 L:317 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:6 M:112 Repeated in SeqNo=6 M:254 Repeated in SeqNo=6 L:374 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:5 L:375 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6 L:375 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:6 L:375 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:6
L:375 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:375 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:6
L:375 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:6
L:375 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:376 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:9
L:378 M:325 E: No of Seq differe (211) IPMCTH INDUCTION FOUNDATION SEQ. L:378 M:252 E: No. of Seq. differs, <211>LENGTH:Input:1090 Found:1107 SEQ:6

BF TECH CENTER 1800 COR